

Rezension: Stephan Büttner, Hans-Christoph Hobohm, Lars Müller (Hg.), 2011: Handbuch Forschungsdatenmanagement

Müller, Karl H.

Veröffentlichungsversion / Published Version

Rezension / review

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Müller, K. H. (2012). Rezension: Stephan Büttner, Hans-Christoph Hobohm, Lars Müller (Hg.), 2011: Handbuch Forschungsdatenmanagement. [Rezension des Buches *Handbuch Forschungsdatenmanagement*, von S. Büttner, H.-C. Hobohm, & L. Müller]. *Methoden, Daten, Analysen (mda)*, 6(1), 53-57. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-314539>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

besseres Verständnis sollten bei zukünftigen Befragungen diese situativen Umstände (Zeitpunkt, Haushaltskontexte, Stimmung, Wetter, etc.) tatsächlich erhoben werden.

Diese Arbeit zeigt, dass für die Erforschung von Nonresponse und Attrition noch ein erhebliches Nachholpotential besteht. Einige äußerst interessante Anregungen werden von der Autorin am Beispiel einer Onlineerhebung von Studierenden gegeben. Es wäre zu begrüßen, wenn diese Anregungen in weiterführenden Arbeiten aufgenommen würden.

Literatur

- AAPOR (The American Association for Public Opinion Research) 2011: Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition (6.3.2012).
- Ehling, M., U. Rendtel et al., 2003: CHINTEX (The Change from Input Harmonisation to Ex-post Harmonisation in National Samples of the European Community Household Panel – Implications on Data Quality) – Synopsis. <http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Wissenschaftsforum/Chintex/ResearchResults/Downloads/Synopsis.psml> (5.3.2012).
- Groves, R. und E. Peytcheva, 2008: The Impact of Non-Response Rates on Non-Response Bias: A Meta-Analysis. *Public Opinion Quarterly* 72 (2): 167–189.
- Lynn, P. und P. Clarke, 2002: Separating Refusal Bias and Non-Contact Bias: Evidence from UK National Surveys. *The Statistician* 51 (3): 319–333.
- Røed K., 2006: Longitudinal Administrative Registers in Economic Research – A Norwegian Experience. Unpublished Presentation for the Conference on Longitudinal Surveys in International Perspective, Montreal, 25–27 January 2006.
- Schouten, B., F. Cobben und J. Bethlehem, 2009: Indicators for the Representativeness of Survey Response. *Survey Methodology* 35 (1): 101–113.
- Voorpostel, M. und O. Lipps, 2011: Attrition in the Swiss Household Panel: Is Change Associated With Later Drop-Out? *Journal of Official Statistics* 27 (2): 301–318.

OLIVER LIPPS, LAUSANNE



STEPHAN BÜTTNER,
HANS-CHRISTOPH
HOB OHM, LARS
MÜLLER (Hg.),
2011: Handbuch
Forschungsdaten-
management.
Bock + Herchen
Verlag, Bad Honnef,
ISBN 978-3-88347-
283-6, 24,90 EUR.
Online Version:
<http://www.forschungsdatenmanagement.de/>

Man trifft immer wieder auf die teilweise paradoxe Situation, dass die Vorteile und Stärken wissenschaftlicher Methoden mit nicht erbringbaren Leistungen legitimiert werden oder dass zentrale Aspekte von Forschungsprozessen kaum oder nur ungenügend thematisiert werden, obwohl auf sie beträchtliche Teile des wissenschaftlichen Arbeitsprozesses entfallen. Beispielsweise gehört es zum Kern des wissenschaftlichen Selbstverständnisses, intersubjektiv zugängliche und reproduzierbare Ergebnisse zu produzieren. Aber ähnlich den Vergessenskurven von Hermann Ebbinghaus gehen quer über die verschiedenen Wissenschaftsfelder sehr schnell die meisten relevanten Datenbestände verloren. Wissenschaftliche Resultate erweisen sich daher gerade *nicht* als intersubjektiv zugänglich oder als reproduzierbar. Ein anderer Punkt betrifft den Nutzen wie den Vorteil von Sekundäranalysen sowie das mit der Zeit wachsende Ergebnispotential solcher Sekundäranalysen. Wegen der vielfach unzugänglichen Daten wird auf dieses Potential weitgehend verzichtet und stattdessen auf jeweils neue Datenerhebungen gesetzt, denen jeweils weitere neue Datenerfassungen folgen. Die akkumulierten Datenbestände aus der Vergangenheit, obwohl sie aus kontextuellen oder aus langfristigen Gesichtspunkten eine starke analytische Bereicherung darstellen würden, bleiben in ihrer überwiegenden Mehrzahl

ungenutzt. Diese beiden Punkte – Datenverluste und intersubjektive Sperren sowie das steigende, aber nicht bearbeitete Potenzial von Sekundäranalysen – gehören gegenwärtig nachgerade zu den blinden Flecken des Wissenschaftssystems, die anscheinend toleriert und hingenommen werden, weil sie, obschon unmittelbar vorhanden, gar nicht erst gesehen werden.

Es gäbe aber auch eine andere Näherung an das Generalthema dieses Buches, das mit bestehenden Anreizsystemen und Leitindikatoren zu tun hat. Auf der individuellen Ebene von Forscherinnen oder Forschern hat sich in den letzten Jahrzehnten ein schmales Set an Schlüsselindikatoren herausgebildet, das einen gewichtigen Teil von Karriereverläufen bestimmt und in dem die wissenschaftliche Produktion vorzugsweise in international akklamierten Zeitschriften, die Durchführung von Forschungsprojekten oder auch die Teilnahme an Konferenzen und Tagungen eine dominante Rolle einnehmen. Genau dieses Set an Schlüsselindikatoren wird auch dazu verwendet, wenn es um Evaluationen von einzelnen Instituten, Fächern oder Universitäten geht. Normalerweise wird es evaluativ als irrelevant erachtet, wie aktiv oder inaktiv sich einzelne WissenschaftlerInnen, Institute oder Forschungsorganisationen im Bereich ihrer Datenorganisation verhalten, ob solche Daten sinnvoll dokumentiert und für den jeweils internen Gebrauch vorhanden sind, ob derartige Daten auch extern zugänglich sind oder ob ein entsprechender Transfer an institutionalisierte Datenarchive erfolgt. Diese individuelle oder institutionelle Forschungsperformanz auf dem Gebiet des Datenmanagements wird evaluativ nahezu völlig ausgeblendet.¹

Mit diesen beiden Annäherungen – einmal aus den Paradoxien von gegenwärtigen Forschungsprozessen und einmal aus der momentanen Marginalisierung eines an sich durchaus zentralen Bereichs in der Forschungsperformanz – kann ein weiterer Kontext aufgespannt werden, in dem das vorliegende Handbuch seine besonderen Wirkungen und Effekte entfalten sollte. Denn mit dem vorliegenden „Handbuch Forschungsdatenmanagement“ wird verdienstvollerweise der Fokus auf unterschiedliche Aspekte in der Entstehung, der Systematisierung und der langfristigen Archivierung von Forschungsdaten gelenkt. Als Besonderheit sei vermerkt, dass die Herausgeber des Handbuches allesamt im Bereich der Informationswissenschaften angesiedelt sind und die AutorInnen der einzelnen Beiträge einer heterogenen Mischung aus Informationswissenschaften, e-Science, einzelnen Fachwissenschaften entstammen.

Damit kann einleitend gleich auf den größten Vorteil dieses Handbuchs verwiesen werden, der einfach darin besteht, ein bislang marginalisiertes Phänomen zu einem Kern- und Schlüsselthema befördert zu haben. Das Forschungsdatenmanagement betrifft im Prinzip alle wissenschaftlichen Felder von der Astronomie hin zur Zoologie – aber erst durch die Durchdringung der Forschungsprozesse mit Informations- und Kommunikationstechnologie (IKT) und deren Vernetzungspotenziale avanciert dieses Thema zu einer gemeinsamen Herausforderung quer durch die einzelnen Wissenschaftsdisziplinen. Forschungspragmatisch steht gegenwärtig ein Paradigmenwechsel an, der das Forschungsdatenmanagement von einer unzureichend durchgeführten *ex post* Aktivität am Ende von Forschungsarbeiten hin zu einem ebenso kontinuierlichen wie notwendigen Prozess schon *ex ante* transformiert, womit nicht nur das Forschungsdatenmanagement von seiner implizit-hintergründigen in eine explizit-vordergründige Rolle wechselt, sondern sich auch die Komposition von Forschungsteams signifikant verändert.

1 Besonders verquer mutet es an, wenn im Rahmen von notorisch gewordenen Evaluationen Forschungsinfrastruktureinrichtungen ihrerseits primär nach den üblichen Kriterien von Publikationen und Forschungsaktivitäten bewertet werden.

Und damit kann gleich ein zweiter Vorteil dieses Handbuches benannt werden, das sich nicht auf ein einzelnes Feld – beispielsweise auf die Sozial- und Wirtschaftswissenschaften – beschränkt, sondern diese Thematik für den gesamten Wissenschaftshaushalt beleuchtet. Die vielfältigen Verweise auf unterschiedliche Problemlagen des Forschungsdatenmanagements in Bereichen wie der Klimaforschung, der Geologie, den Sozial- und Wirtschaftswissenschaften oder den *Life-Sciences* beziehungsweise die neuen Herausforderungen für Informationswissenschaftler als *data curators* oder *data librarians* befördern dieses Handbuch in die Kategorie der Pflichtlektüre für Forscherinnen oder Forscher aus den unterschiedlichsten Gebieten und zu einem fixen Bezugspunkt für die Bereiche der Wissenschaftspolitik und der Wissenschaftsförderung.

Nach dieser anfänglichen Eloge kann näher auf die Inhalte dieses Handbuchs überschwenkt werden, die sich in drei großen Themengruppen bündeln, nämlich in

- allgemeine Fragen des Forschungsdatenmanagements,
- den Lebenszyklus von Forschungsdaten,
- die institutionalisierten Formen der Datenarchivierung.

Im ersten Teil wird mit drei Übersichtsartikeln zum Bereich „Research Data Management“, zum „Lebenszyklus von Forschungsdaten“ sowie zum „Datenmanagement und Data Sharing“ in den Gesamtbereich eingeführt. Dabei wird im ersten Artikel nicht nur die Heterogenität des Datenbegriffs in den unterschiedlichen Wissenschaftsfeldern beleuchtet, sondern auch die besondere Rolle, welche die Informationswissenschaften im Ausbau von Forschungsdaten und Forschungsinfrastrukturen spielen können. Der zweite Artikel stellt die beiden grundlegenden Modelle des Forschungsdatenmanagements vor, nämlich das *Curation Life Cycle*-Modell des britischen *Digital Curation Centre* (DCC) sowie das australische *Data Curation Continuum*, das den Lebenszyklus von Forschungsdaten von ihrer Entstehung

bis hin zu ihrer langfristigen digitalen Verfügbarkeit beschreibt. Der dritte Artikel rekurriert auf die besonderen Erfahrungen mit Datenmanagement und Data Sharing im Bereich der Sozial- und Wirtschaftswissenschaften, wobei neben den fachspezifischen Datenarchiven auch dezentrale Forschungsdatenzentren als zukünftige Option in den Raum gestellt werden.

Der zweite Teil des Handbuchs umfasst dann die Hälfte aller Beiträge und widmet sich einzelnen Aspekten des Forschungsdatenmanagements: den „Data Policies“, den rechtlichen Problemen der elektronischen Langzeitarchivierung, den „Datenmanagementpläne(n)“, der Frage von „Metadaten und Standards“, den „Forschungsdaten-Repositorien“, der „Langzeiterhaltung digitaler Forschungsdaten“, den „Systeme(n) und Systemarchitekturen für das Datenmanagement“ sowie der „Datenanalyse und -visualisierung“.

Der erste Artikel im zweiten Teil vermittelt einen nützlichen Überblick zum Stellenwert und zu den Dimensionen von Daten-Politiken in den unterschiedlichsten Wissenschaftskontexten, wobei auch die notorischen Probleme solcher Daten-Politiken angesprochen werden. Der zweite Artikel versammelt die rechtlichen Probleme im Forschungsdatenmanagement – und hier insbesondere das Problem des Datenschutzes wie auch des Urheberrechts. Der dritte Abschnitt widmet sich der Organisation wie den Strukturen von Datenmanagementplänen vor allem im Hinblick auf ein nachhaltiges Datenmanagement. Mit dem vierten Artikel wird auf die Problematik von einheitlichen Standards von Metadaten verwiesen, die eine unentbehrliche Rolle für den Bereich der Datensuche und damit auch für das Data-Sharing spielen. Der fünfte Abschnitt führt vor allem in die technischen Unterbauten von Repositorien ein, in denen auch Daten versammelt und für die Weiternutzung zur Verfügung gestellt werden. Der sechste Artikel schließt mit einem überaus nützlichen Überblick zu Konzepten, Methoden und Werkzeugen der Langzeit-

archivierung an. Mit dem siebenten Artikel wird ein umfassender Einblick in die Systeme und die Systemarchitekturen im Forschungsdatenmanagement offeriert, wobei dabei der Grid-Technologie ein besonderer Stellenwert eingeräumt wird. Und schließlich wurde auch ein Artikel über die neuen Potentiale der Datenvisualisierung inkludiert, der indirekt die Datenvisualisierung als ein genuines und eigenständiges Methodenfeld speziell innerhalb der entstehenden e-Science Infrastrukturen anspricht (vgl. dazu auch Müller/Reautschnig, 2010; 2011).

Und schließlich behandelt der dritte Teil stärker institutionelle wie ausbildungsrelevante Aspekte der Langzeitarchivierung – und dies in Form eines Beitrags zu „Data Curation Services“, zu „Strategien zur Veröffentlichung von Forschungsdaten“, zur „Archivierung von Forschungsdaten“ sowie zum Bereich „Informationswissenschaftler im Forschungsdatenmanagement“.

Der erste Artikel vermittelt eine instruktive Übersicht zu den gegenwärtigen Transformationen hin zu langfristig angelegten *Data Curation Services*, welche die dauerhafte Nutzung von Datenbeständen garantieren sollen. Der zweite Abschnitt stellt interessante Publikationsmodelle für den Datenbereich vor, durch die der momentane Status der Datenerstellung und Datenbereitstellung verschieden aufgewertet werden könnte. Der dritte Artikel führt in ein enorm wichtiges Gebiet, in dem es um die semantischen Vernetzungen von Forschungsdaten mit Hilfe unterschiedlicher Ontologiesprachen geht. Mit dem vierten Abschnitt werden nochmals die besonderen Probleme der Langzeitarchivierung im Spannungsfeld von nicht verwendbarem Datenmüll und notwendiger Datenpflege thematisiert. Und schließlich wird die besondere Rolle der Informationswissenschaften herausgearbeitet, denen im Zuge der wachsenden Wichtigkeit des Forschungsdatenmanagements neue Aufgaben, Leistungen und curriculare Ausbildungswege zuwachsen.

Auch ein sehr aufschlussreiches Buch bietet keinen Schutz davor, dass ein Rezensent trotz alledem einige formale sowie auch inhaltliche Mängel zur Sprache bringt.

Auf einer rein formalen Ebene hätte man sich gerade bei einem so interessanten Buch einen liebevolleren und auch selbstreferentiell stimmigeren Umgang mit den Inhalten gewünscht. Beispielsweise sind, um mit einem kleinen Punkt zu beginnen, die Grafiken auf den Seiten 27 und 30 kaum beziehungsweise nicht lesbar. Weiter fehlt am Schluss ein entsprechender Sachindex – ein Autorenindex könnte ja angesichts der Diversität der Themenpalette aus guten Gründen entfallen. Auch wurde auf eine gemeinsame Bibliografie am Ende des Bandes verzichtet – und man findet die bibliografischen Hinweise nur jeweils am Ende der einzelnen Kapitel. Und schließlich wäre es gerade bei der Wichtigkeit und Neuartigkeit des Themas von Belang gewesen, hätte man am Schluss eine Kompilation wichtiger Schlüsselbegriffe aus dem Bereich Forschungsdatenmanagement vorgenommen und deren Bedeutungen zusammenfassend beschrieben.

Inhaltlich hätte man sich zumindest einen vierten Teil erhofft, in dem – symmetrisch zur Einleitung – in zwei oder drei Artikeln die gegenwärtigen wissenschaftspolitischen Implikationen thematisiert wie auch die zukünftigen Herausforderungen des Forschungsdatenmanagements deutlicher herausgearbeitet worden wären. Diese Herausforderungen sind vor allem deswegen so enorm, weil sich in einzelnen Wissenschaftsfeldern – speziell in den Sozial- und Wirtschaftswissenschaften – langsam ein stärkerer Wechsel hin zu prozessproduzierten Daten vollziehen sollte. Solche prozessproduzierten Daten im Bereich Konsum, Informationsverhalten oder anderer Online-Aktivitäten stellen völlig neuartige Probleme in der Datenkonstruktion, der Datenanalyse wie auch in der Langzeitarchivierung. Weiter – um einen weiteren zukunftsrelevanten Punkt zu benennen – erweisen sich Datencluster aus

unterschiedlichen disziplinären Feldern im Prinzip als kombinierbar. Solche Kombinationen scheitern aber gegenwärtig an unterschiedlichen Datenformaten und Standards sowie an den institutionalisierten Grenzen von Datenarchivierungen. Diese technologischen, semantischen und institutionellen Probleme der Datenkombination quer über etablierte Disziplinfelder hätten es durchaus verdient, deutlicher thematisiert zu werden. Und schließlich bedeutet ein stetig begleitendes Forschungsdatenmanagement auch die Schaffung entsprechender Ausbildungs- und Berufsfelder, für die neben neuen Kompetenzen und Ausbildungen im Bereich der Informationswissenschaften auch spezielle disziplinspezifische Kompetenzen und Curricula notwendig werden.

Aber diese formalen wie inhaltlichen Punkte eines kleineren Missmanagements im Bereich des Forschungsdatenmanagements wiegen vergleichsweise gering gegenüber den angebotenen Inhalten dieses Buches, das wegen der immer größeren Wichtigkeit des Themas und wegen der Breite der angesprochenen Inhalte zu einem *Instant*-Klassiker avancieren müsste. Und es bleibt zu hoffen, dass durch dieses Werk auch entsprechende wissenschaftspolitische Diskussionen zum Stellenwert des Forschungsdatenmanagements in Gang kommen, die bis hin zu institutionellen, regionalen, nationalen oder europäischen Datenpolitiken reichen sollten.

Literatur

- Müller, K. H. und A. Reautschnig, 2010: Die visuelle Datenanalyse (VDA) in der vergleichenden sozialwissenschaftlichen Forschung. S. 236-250 in: Hug, T. und A. Kriwak (Hg.): Visuelle Kompetenz. Beiträge des interfakultären Forums Innsbruck Media Studies. Innsbruck: Innsbruck University Press.
- Müller, K. H. und A. Reautschnig, 2011: From Otto Neureath's ISOTYPE to Multiple Worlds of Media. S. 185-214 in: Heinrich, R., E. Nemeth, W. Pircher und D. Wagner (Hg.), Image and Imaging in Philosophy, Science and the Arts, Bd. 2. Frankfurt:ontos Verlag.

KARL H. MÜLLER, WIEN